

(19) World Intellectual Property
Organization
International Bureau



542208

(43) International Publication Date
29 July 2004 (29.07.2004)

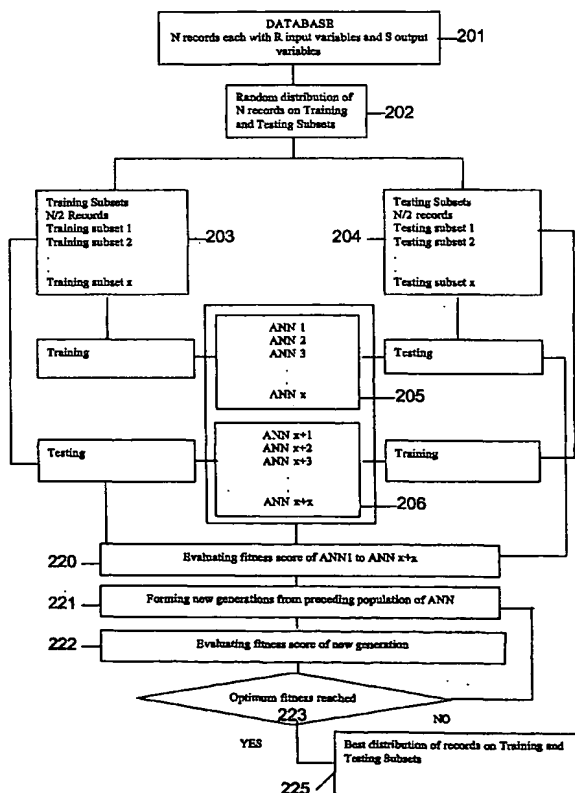
PCT

(10) International Publication Number
WO 2004/063831 A2

- (51) International Patent Classification⁷: G06F (74) Agent: KARAGHIOSOFF, Giorgio, A.; 27 V. Pecorile, I-17015 Celle Ligure (SV) (IT).
- (21) International Application Number: PCT/EP2004/000157 (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) International Filing Date: 13 January 2004 (13.01.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/440,210 15 January 2003 (15.01.2003) US
- (71) Applicant (for all designated States except US): BRACCO IMAGING S.P.A. [IT/IT]; 50 v. Folli, I-20134 Milano (IT).
- (72) Inventor; and
- (73) Inventor/Applicant (for US only): BUSCEMA, Massimo [IT/IT]; 47 v. N. Coviello, I-00163 Roma (IT).
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK,

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR OPTIMIZATION OF A DATABASE FOR THE TRAINING AND TESTING OF PREDICTION ALGORITHMS



(57) Abstract: A system and method are provided for the training and testing of prediction algorithms. According to an exemplary embodiment of the invention the method generates optimum training, testing and/or validation data sets from a common general database by applying a genetic algorithm to populations of testing and training subsets used in connection with a given prediction algorithm. In exemplary embodiments the prediction algorithm operated upon is an artificial neural network. As well, in preferred exemplary embodiments, the most predictive independent variables of the records of the common database are automatically selected in a pre-processing phase. Such preprocessing phase applies a genetic algorithm to populations of prediction algorithms which vary as to number and content of input variables, where the prediction algorithms representing the selections of input variables which have the best testing performances and the minimum input variables are promoted for the processing of the new generations according to a defined selection algorithm.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *without international search report and to be republished upon receipt of that report*